

AAB03792

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DR

PT

XX

XX

XX

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	5
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---

Db 1 MSDHGDVSLPPQDRVRILSQLGSAVELNEDI PPRRYRSGVEI I RMASVYSEEGNIEHAF 60

DR WPI; 2000-387743/33.
 DR N-PSDB; AAA59852.
 XX
 PT Protein AMSH being signal transducer interacting with SH3 domain of
 PT cytokine signal transducer STAM, for diagnosis and development of drugs
 PT of treating diseases e.g. severe infection, cancers and autoimmune
 PT disorders.
 XX
 PS Claim 1; Page 12-14; 22pp; Japanese.
 XX
 CC This sequence represents the human AMSH protein. The AMSH protein is
 CC involved in signal transduction, and interacts with the SH3 domain of
 CC cytokine signal transducer STAM. The invention includes human and murine
 CC AMSH protein and gene sequences (hAMSH and mAMSH), and also includes
 CC antibodies directed against hAMSH and mAMSH. The AMSH protein can be used
 CC in the diagnosis and development of drugs for treating diseases due to
 CC the dysfunction of cytokine signal transduction e.g. severe infection,
 CC cancers and autoimmune disorders
 XX
 SQ Sequence 424 AA;

Query Match 85.2%; Score 1904; DB 3; Length 424;
 Best Local Similarity 83.5%; Pred. No. 3.4e-150;
 Matches 354; Conservative 35; Mismatches 35; Indels 0; Gaps 0;

QY	1	MSDHGDVSLPPQDRVRILSQLGSAVELNEDI PPRRYRSGVEI I RMASVYSEEGNIEHAF	60
		: : : : : : :	
Db	1	MSDHGDVSLPPEDRVRALSQLGSAVEVNEDI PPRRYFRSGVEI I RMASIYSEEGNIEHAF	60
QY	61	ILYNKYITLFI EKLPKHRDYKSAI I PEKKDAVKKLKSVAFPKAEELKTELLRRYTKEYEQ	120
		: : : : : : :	
Db	61	ILYNKYITLFI EKLPKHRDYKSAVI PEKKDVTVKKLKEIAFPKAEELKAEELKRYTKEYTE	120
QY	121	YKERKKKEEEELARNIAIQQELEKEKQORVAQQKQKQLEQEQFHAFEEMIQRQLEKERLK	180
		: : : : : : :	
Db	121	YNEEKKKEAEELARNMAIQQELEKEKQORVAQQKQQQLEQEQFHAFEEMI RNQLEKERLK	180
QY	181	IVQEF GKVDPGPCGPLLPDLEKPCVDVAPSSPFSPQTPTDCNTGMRPAKPPVDRSLKPG	240
		: : : : : : : : :	
Db	181	IVQEF GKVDPGLGGLVDPDLEKPSLDVFPTLT VSSIQPSDCHTTVRPAKPPVDRSLKPG	240
QY	241	ALSVIENVPTIEGLRHI VVPRNLCSEFLQLASANTAKGIETCGVLCGKLMRNEFTITHVL	300
		: : : : : : : : : : : : :	
Db	241	ALSNSESIPTIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMRNEFTITHVL	300
QY	301	IPRQNGGPDYCHTENEIEIFFMQDDLGLLTLGWIHTHTPTQTAF LSSVDLHTHCSYQMMLP	360
		: : : : : : : : : : : : :	
Db	301	IPKQSAGSDYCNTENEEELFLIQDQQGLITLGWIHTHTPTQTAF LSSVDLHTHCSYQMMLP	360
QY	361	ESIAIVCSPKFQETGFFKLT DYGLQEISTCRQKGFHPHGRDPPLFCDCSHVTVKDRIVTI	420
		: : : : : : : :	
Db	361	ESVAIVCSPKFQETGFFKLT DHGLEEISSCRQKGFHPHSDPPLFCSCSHVTVVDRAVTI	420
QY	421	TDLR	424
Db	421	TDLR	424

RESULT 3

AAB62193

ID AAB62193 standard; protein; 424 AA.

XX

AC AAB62193;

XX

DT 11-JUN-2001 (first entry)

XX

DE Human SAP1/AMSH amino acid sequence.

XX

KW Smad associating polypeptide; SAP; cell cycle; TGF-beta; SAP1/AMSH;

KW signal transduction; phosphorylation; gene therapy; human.

XX

OS Homo sapiens.

XX

PN WO200121794-A2.

XX

PD 29-MAR-2001.

XX

PF 20-SEP-2000; 2000WO-US025790.

XX

PR 20-SEP-1999; 99US-0154846P.

XX

PA (LUDW-) LUDWIG INST CANCER RES.

XX

PI Itoh F, Itoh S, Heldin C, Ten-Dijke P;

XX

DR WPI; 2001-257987/26.

DR N-PSDB; AAF57478.

XX

PT Novel Smad associating proteins useful for regulating cell cycle in a
PT mammalian cell, and for identifying lead compounds for pharmacological
PT agents.

XX

PS Claim 14; Page 53-55; 90pp; English.

XX

CC The invention relates to Smad associating polypeptides (SAP) and nucleic
CC acids encoding the SAPs. The SAPs can be expressed by standard
CC recombinant methodology. The SAP proteins and polynucleotides are useful
CC for regulating the cell cycle in a mammalian cell, and for modulating TGF
CC -beta superfamily signal transduction in a mammalian cell. SAP
CC polypeptide are useful for testing Smad binding, for testing
CC phosphorylation, for generating antibodies, and as components of an
CC immunoassay. The SAP polynucleotides are useful in gene therapy. The
CC present sequence represents the amino acid sequence of human SAP1/AMSH
CC that can be used in modulating TGF-beta superfamily signal transduction
CC in a mammalian cell

XX

SQ Sequence 424 AA;

Query Match 85.2%; Score 1904; DB 4; Length 424;

Best Local Similarity 83.5%; Pred. No. 3.4e-150;

Matches 354; Conservative 35; Mismatches 35; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPQDRVRILSQLGSAVELNEDIPPRRYRSGVEIIRMASVYSEEGNIEHAF 60
|||||||:||||:|||||||:|||||||:|||||||:|||||||

6,673,570

Db 301 IPKQSAGSDYCNTEEEEEELFIQDQQLITLGWIHTHTPTQTAFLLSSVDLHTHCSYQMMLP 360

QY 361 ESIAIVCSPKFQETGFFKLTDYGLQEISTCRQKGFHHPGRDPPLFCDCSHVTVKDRIVTI 420
 ||:||||||||||||||||||:||||:||||||||||:||||| || |||

Db 361 ESVAIVCSPKFQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420

QY 421 TDLR 424
 ||||

Db 421 TDLR 424

RESULT 5

ABB81014

ID ABB81014 standard; peptide; 424 AA.

XX

AC ABB81014;

XX

DT 21-OCT-2002 (first entry)

XX

DE Human AMSH protein sequence.

XX

KW JAM; Jab1-associated metalloenzyme motif; peptidase; cytostatic;
 KW antimicrobial; antiinflammatory; antiasthmatic; vasotropic; AMSH;
 KW neuroprotective; antirheumatic; antiarthritic; antipsoriatic; human.

XX

OS Homo sapiens.

XX

PN WO200255536-A2.

XX

PD 18-JUL-2002.

XX

PF 14-JAN-2002; 2002WO-US000976.

XX

PR 12-JAN-2001; 2001US-0261314P.

PR 14-SEP-2001; 2001US-0322030P.

PR 14-SEP-2001; 2001US-0322322P.

XX

PA (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX

PI Cope G, Verma R, Aravind L, Koonin EV, Deshaies R;

XX

DR WPI; 2002-590657/63.

XX

PT New polypeptide, useful for screening for compounds for treating
 PT neoplastic growth, angiogenesis, chronic inflammation, asthma, ischemia,
 PT multiple sclerosis, rheumatoid arthritis or psoriasis, comprises Jab1-
 PT associated metalloenzyme motif.

XX

PS Disclosure; Fig 2; 35pp; English.

XX

CC The invention relates to an isolated (crystalline) polypeptide (I)
 CC comprising a Jab1-associated metalloenzyme motif (JAM) domain, where the
 CC JAM domain is not adjacent to an amino acid sequence that is naturally
 CC adjacent to the domain. (I) has peptidase activity and can be used for
 CC modulation of neddylation/de-neddylation; ubiquitination/de-
 CC ubiquitination and as COP9/signalsome, 26S proteasome, AMSH, AMSH1 or
 CC AMSH2 inhibitors. Inhibitor for (I) that will form a bond with the JAM

2003/0166243

CC domain and modulators of (I) are useful for treating neoplastic growth,
 CC angiogenesis, infection, chronic inflammation, asthma, ischaemia and
 CC reperfusion injury, multiple sclerosis, rheumatoid arthritis or
 CC psoriasis. The present sequence represents a human AMSH protein
 XX

SQ Sequence 424 AA;

Query Match 85.2%; Score 1904; DB 5; Length 424;
 Best Local Similarity 83.5%; Pred. No. 3.4e-150;
 Matches 354; Conservative 35; Mismatches 35; Indels 0; Gaps 0;

Qy	1	MSDHGDVSLPPQDRVRILSQLGSAVELNEDIIPRRYYRSGVEIIRMASVYSEEGNIEHAF	60
Db	1	MSDHGDVSLPPEDRVRLSQLGSAVEVNEDIIPRRYFRSGVEIIRMASIYSEEGNIEHAF	60
Qy	61	ILYNKYITLFIKLPKHRDYKSAI IPEKKDAVKKLKSVAFPKAEELKTELLRRYTKEYEQ	120
Db	61	ILYNKYITLFIKLPKHRDYKSAVIPEKKDVTVKKLKEIAFPKAEELKAELLKRYTKEYTE	120
Qy	121	YKERKKKEEEEELARNIAIQQELEKEKQORVAQQKQKQLEQEQFHAFEEIMQRQLEKERLK	180
Db	121	YNEEKKKAEELARNMAIQQELEKEKQORVAQQKQKQLEQEQFHAFEEIMRNQLEKERLK	180
Qy	181	IVQEFGKVDPGPCGPLLPDLEKPCVDVAPSSPFSPTQTPDCNTGMRPAKPPVVDRLSKPG	240
Db	181	IVQEFGKVDPLGGPLVPDLEKPSLDVFPPLTVSSIQPSDCHTTVRPAKPPVVDRLSKPG	240
Qy	241	ALSVIENVPTIEGLRHI VVPRNLCSEFLQLASANTAKGIETCGVLCGKLMRNEFTITHVL	300
		: :	
Db	241	ALSNSESIPTIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMRNEFTITHVL	300
Qy	301	IPRQNGGPDYCHTENESEEIFFMQDDLGLLTLGWIHTHTPTQTAFSSVDLHTHCSYQMMLP	360
		: :	
Db	301	IPKQSAGSDYCNTENEEELFLIQDQOGLITLGWIHTHTPTQTAFSSVDLHTHCSYQMMLP	360
Qy	361	ESIAIVCSPKFQETGFFKLTDYGLQEISTCRQKGFHHPGRDPPLFCDCSHVTVKDRIVTI	420
		:	
Db	361	ESVAIVCSPKFQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI	420
Qy	421	TDLR	424
Db	421	TDLR	424

RESULT 6

ABW01032

ID ABW01032 standard; protein; 424 AA.

XX

AC ABW01032;

XX

DT 15-JAN-2004 (first entry)

XX

DE Human AMSH protein.

XX

KW Jab1-associated metalloenzyme motif; JAMM domain; isopeptidase; asthma;
 KW drug design; proliferative disorder; inflammation; autoimmune disorder;
 KW rheumatoid arthritis; neuroprotective; COP9 signalsome complex; CSN;

KW E2 ubiquitin conjugating enzyme; degradation pathway; multiple sclerosis;
 KW angiogenesis; psoriasis; cytostatic; ischaemia; vasotropic; human.
 XX
 OS Homo sapiens.
 XX
 PN US2003153097-A1.
 XX
 PD 14-AUG-2003.
 XX
 PF 09-JAN-2003; 2003US-00340578.
 XX
 PR 12-JAN-2001; 2001US-0261314P.
 PR 14-SEP-2001; 2001US-0322030P.
 PR 14-SEP-2001; 2001US-0322322P.
 PR 14-JAN-2002; 2002US-00046961.
 PR 14-JAN-2002; 2002US-00047253.
 PR 06-FEB-2002; 2002US-0355334P.
 XX
 PA (DESH/) DESHAIES R J.
 PA (COPE/) COPE G.
 PA (VERM/) VERMA R.
 PA (AMBR/) AMBROGGIO X I.
 XX
 PI Deshaies RJ, Cope G, Verma R, Ambroggio XI;
 XX
 DR WPI; 2003-811195/76.
 XX
 PT Isolated COP9 signalsome isopeptidase comprising a mutant JAMM domain,
 PT useful for identifying the agents for treating cancer, autoimmune disease
 PT by modulating the peptidase activity.
 XX
 PS Example 3; Fig 1; Opp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising a mutant JAMM
 CC (Jabl-associated metalloenzyme motif) domain. JAMM domain is useful for
 CC identifying an inhibitor of an isopeptidase activity of a polypeptide by
 CC rational drug design. JAMM is useful for treating proliferative disorder,
 CC inflammatory disorder or an autoimmune disorder. The cell proliferative
 CC disorder is a neoplastic growth or psoriasis. The inflammatory disorder
 CC is an acute infection or a chronic inflammatory disorder. The autoimmune
 CC disorder is multiple sclerosis or rheumatoid arthritis. The conditions is
 CC angiogenesis, asthma or ischaemia and reperfusion injury. The polypeptide
 CC comprising the JAMM domain is JABI. The polypeptide comprising the JAMM
 CC domain comprises a COP9 signalsome (CSN) complex. The method further
 CC involves administering to the subject at least a second agent that
 CC modulates a component of a protein degradation pathway comprising the
 CC polypeptide comprising the JAMM domain. The second agent reduces or
 CC inhibits the activity of the component of the protein degradation
 CC pathway. The component of the protein degradation pathway comprises an E2
 CC ubiquitin conjugating enzyme or an E2 ubiquitin ligase. The E2 ubiquitin
 CC conjugating enzyme is Cdc34 or a Ubc4/5 family member. The E3 ubiquitin
 CC ligase comprises Skp1, Cull/Cdc53, F-box protein, or their combination.
 CC The E3 ubiquitin ligase comprises SCF ubiquitin ligase. The present
 CC sequence is human AMSH protein used in the exemplification of the
 CC invention
 XX
 SQ Sequence 424 AA;

Query Match 85.2%; Score 1904; DB 7; Length 424;
Best Local Similarity 83.5%; Pred. No. 3.4e-150;
Matches 354; Conservative 35; Mismatches 35; Indels 0; Gaps 0;

```
Qy      1 MSDHGDVSLPPQDRVRILSQLGSAVELNEDIPRRYYRSGVEIIRMASVYSEEGNIEHAF 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1 MSDHGDVSLPPEDRVRLSQLGSAVEVNEDIPRRYFRSGVEIIRMASIYSEEGNIEHAF 60

Qy     61 ILYNKYITLFIKLPKHRDYKSAI IPEKKDAVKKLKSVAFPKAEELKTELLRRYTKEYEQ 120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     61 ILYNKYITLFIKLPKHRDYKSAVIPEKKDTVKKLKEIAFPKAEELKAELLKRYTKEYTE 120

Qy    121 YKERKKKEEEELARNIAIQQELEKEKQORVAQQKQKQLEQEQFHAFEEMIQRQLEKERLK 180
        || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    121 YNEEKKEAEELARNMAIQQELEKEKQORVAQQKQQQLEQEQFHAFEEIIRNQLEKERLK 180

Qy    181 IVQEFQKVDPGPCGPLLPDLEKPCVDVAPSSPFSPTQTPDCNTGMRPAKPPVVDRSLKPG 240
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    181 IVQEFQKVDPLGGPLVPDLEKPSLDVFPPTLTVSSIQPSDCHTTVRPAKPPVVDRSLKPG 240

Qy    241 ALSVIENVPTIEGLRHI VVPRNLCSEFLQLASANTAKGIETCGVLCGKLMRNEFTITHVL 300
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    241 ALSNSES IPTIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMRNEFTITHVL 300

Qy    301 IPRQNGGPDYCHTENEIEIFFMQDDLGLLTLGWIHTHTPTQTAFSSVDLHTHCSYQMMLP 360
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    301 IPKQSAGSDYCNTENEEELFLIQDQQGLITLGWIHTHTPTQTAFSSVDLHTHCSYQMMLP 360

Qy    361 ESIAIVCSPKFQETGFFKLTDYGLQEISTCRQKGFHHPGRDPPLFCDCSHVTVKDRIVTI 420
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    361 ESVAIVCSPKFQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420

Qy    421 TDLR 424
        ||| |||
Db    421 TDLR 424
```

RESULT 1

Q9CQ26

ID Q9CQ26 PRELIMINARY; PRT; 424 AA.
AC Q9CQ26;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 5330424L14Rik protein (5730422L11Rik protein) (RIKEN cDNA 5730422L11
DE gene) (AMSH) (Associated molecule with the SH3 domain of STAM).
GN STAMPB OR AMSH OR 5330424L14RIK OR 5730422L11RIK OR MAMSH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo, and Pituitary;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzairelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=21570416; PubMed=11713295;
RA Ishii N., Owada Y., Yamada M., Miura S., Murata K., Asao H., Kondo H.,
RA Sugamura K.;
RT "Loss of Neurons in the Hippocampus and Cerebral Cortex of AMSH-
RT Deficient Mice.";
RL Mol. Cell. Biol. 21:8626-8637(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor, and Liver;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK019907; BAB31909.1; -.
DR EMBL; AK017600; BAB30832.1; -.
DR EMBL; BC003497; AAH03497.1; -.
DR EMBL; BC006939; AAH06939.1; -.
DR EMBL; AB010123; BAB78604.1; -.

DR EMBL; BC025111; AAH25111.1; -.
DR MGD; MGI:1917777; Stambp.
DR GO; GO:0006916; P:anti-apoptosis; IMP.
DR InterPro; IPR000555; Peptidase_M67.
DR Pfam; PF01398; Mov34; 1.
DR SMART; SM00232; JAB_MPN; 1.
SQ SEQUENCE 424 AA; 48513 MW; D3A9C2B5F65B7C5E CRC64;

Query Match 100.0%; Score 2236; DB 11; Length 424;
Best Local Similarity 100.0%; Pred. No. 6.8e-149;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MSDHGDVSLPPQDRVRILSQLGSAVELNEDIPRRYYRSGVEIIRMASVYSEEGNIEHAF 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSDHGDVSLPPQDRVRILSQLGSAVELNEDIPRRYYRSGVEIIRMASVYSEEGNIEHAF 60

Qy     61 ILYNKYITLFIKLPKHRDYKSAIIPEKKDAVKKLKSVAFPKAEELKTELLRRYTKEYEQ 120
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ILYNKYITLFIKLPKHRDYKSAIIPEKKDAVKKLKSVAFPKAEELKTELLRRYTKEYEQ 120

Qy    121 YKERKKKEEEEELARNIAIQQELEKEKQORVAQQKQKQLEQEQFHAFEEIMQRQELEKERLK 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 YKERKKKEEEEELARNIAIQQELEKEKQORVAQQKQKQLEQEQFHAFEEIMQRQELEKERLK 180

Qy    181 IVQEF GKVDPGPCGPLL PDLEKPCVDVAPSSPFSPTQTPDCNTGMRPAKPPVDRSLKPG 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IVQEF GKVDPGPCGPLL PDLEKPCVDVAPSSPFSPTQTPDCNTGMRPAKPPVDRSLKPG 240

Qy    241 ALSVIENVPTIEGLRHIVVPRNLCSEFLQLASANTAKGIETCGVLCGKLMRNEFTITHVL 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ALSVIENVPTIEGLRHIVVPRNLCSEFLQLASANTAKGIETCGVLCGKLMRNEFTITHVL 300

Qy    301 IPRQNGGPDYCHTENEIEIFFMQDDLGLLTLGWIHTHTPTQTAF LSSVDLHTHCSYQMMLP 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 IPRQNGGPDYCHTENEIEIFFMQDDLGLLTLGWIHTHTPTQTAF LSSVDLHTHCSYQMMLP 360

Qy    361 ESIAIVCSPKFQETGFFKLT DYGLQEISTCRQKGFHPHGRDPPLFCDCSHVTVKDRIVTI 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ESIAIVCSPKFQETGFFKLT DYGLQEISTCRQKGFHPHGRDPPLFCDCSHVTVKDRIVTI 420

Qy    421 TDLR 424
      ||||
Db    421 TDLR 424
```

RESULT 2

Q8R424

ID Q8R424 PRELIMINARY; PRT; 424 AA.
AC Q8R424;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AMSH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

RESULT 1

E35542

LOCUS E35542 1384 bp DNA linear PAT 31-JAN-2002

DEFINITION Protein AMSH and cDNA thereof.

ACCESSION E35542

VERSION E35542.1 GI:18624563

KEYWORDS JP 2000139469-A/2.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1384)

AUTHORS Sugamura, K. and Tanaka, N.

TITLE Protein AMSH and cDNA thereof

JOURNAL Patent: JP 2000139469-A 2 23-MAY-2000; SCIENCE & TECH AGENCY

COMMENT OS Homo sapiens (human)

PN JP 2000139469-A/2

PD 23-MAY-2000

PF 12-NOV-1998 JP 1998322674

PR

PI KAZUO SUGAMURA, NOBUYUKI TANAKA

PC C12N15/09, C07K14/52, C07K16/24//C12N1/21, C12P21/02, (C12N15/09, C12R1:91),

PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), C12N15/00, (C12N15/00, C12R1:91)

CC

FH Key Location/Qualifiers

FT CDS 56..1327.

FEATURES Location/Qualifiers

source 1..1384

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1384; DB 6; Length 1384;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1  GTGACGTTTCCGGAAGCTCTGACTGTCATCCTTCACGAAAGAACTTATTTGTCCAATGTC 60
      |||
Db      1  GTGACGTTTCCGGAAGCTCTGACTGTCATCCTTCACGAAAGAACTTATTTGTCCAATGTC 60

Qy     61  TGACCATGGGGATGTGAGCCTCCCAACCCCAAGACCGGGTGAGGATTCTGTCCCAACTGG 120
      |||
Db     61  TGACCATGGGGATGTGAGCCTCCCAACCCCAAGACCGGGTGAGGATTCTGTCCCAACTGG 120

Qy    121  GAGTGCAGTTGAGTTAAATGAAGACATTCCACCCCGTCGCTACTACCGCTCCGGTGTGTTGA 180
      |||
Db    121  GAGTGCAGTTGAGTTAAATGAAGACATTCCACCCCGTCGCTACTACCGCTCCGGTGTGTTGA 180

Qy    181  GATCATCCGCATGGCGTCCGTTTACTCGGAAGAAGGCAACATTGAACATGCCTTTATCCT 240
      |||
Db    181  GATCATCCGCATGGCGTCCGTTTACTCGGAAGAAGGCAACATTGAACATGCCTTTATCCT 240

Qy    241  CTACAACAAGTACATCAGCTGTTTATTGAAAACTTCCGAAACACCGAGACTACAAATC 300

```

Db	241		CTACAACAAGTACATCACGCTGTTTATTGAAAAA	CTTCCGAAACACCGAGACTACAAATC	300
Qy	301		AGCTATCATTCCTGAGAAGAAAGATGCTGTCAAGAAATTAAGAGCGTCGCTTTCCCTAA		360
Db	301		AGCTATCATTCCTGAGAAGAAAGATGCTGTCAAGAAATTAAGAGCGTCGCTTTCCCTAA		360
Qy	361		AGCGGAAGAGCTGAAGACAGAGCTCTTGAGAAGATACACCAAAGAATATGAGCAGTATAA		420
Db	361		AGCGGAAGAGCTGAAGACAGAGCTCTTGAGAAGATACACCAAAGAATATGAGCAGTATAA		420
Qy	421		AGAGCGAAAGAAAAAGGAAGAAGAGGAACTTGCCCGAAATATCGCCATCCAGCAAGAGTT		480
Db	421		AGAGCGAAAGAAAAAGGAAGAAGAGGAACTTGCCCGAAATATCGCCATCCAGCAAGAGTT		480
Qy	481		GGAAAAAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATT		540
Db	481		GGAAAAAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATT		540
Qy	541		CCATGCCCTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGT		600
Db	541		CCATGCCCTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGT		600
Qy	601		TCAAGAGTTCGGGAAGGTAGACCCTGGCCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAA		660
Db	601		TCAAGAGTTCGGGAAGGTAGACCCTGGCCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAA		660
Qy	661		GCCTTGTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAA		720
Db	661		GCCTTGTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAA		720
Qy	721		CACAGGCATGAGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTT		780
Db	721		CACAGGCATGAGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTT		780
Qy	781		AAGCGTCATAGAAAAATGTTCCCAACATTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAA		840
Db	781		AAGCGTCATAGAAAAATGTTCCCAACATTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAA		840
Qy	841		TCTGTGCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTG		900
Db	841		TCTGTGCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTG		900
Qy	901		TGGAGTCCTCTGTGGAAAACTGATGAGAAATGAATTACAATCACACATGTTCTCATCCC		960
Db	901		TGGAGTCCTCTGTGGAAAACTGATGAGAAATGAATTACAATCACACATGTTCTCATCCC		960
Qy	961		CAGACAAAATGGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTAT		1020
Db	961		CAGACAAAATGGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTAT		1020
Qy	1021		GCAGGATGACCTTGGA	CTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGC	1080
Db	1021		GCAGGATGACCTTGGA	CTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGC	1080
Qy	1081		CTTTCTGTCCAGTGTGGATCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTC		1140

Db 1081 CTTTCTGTCCAGTGTGGATCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTC 1140

Qy 1141 CATCGCAATCGTCTGTTCCTCCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTA 1200
 |||

Db 1141 CATCGCAATCGTCTGTTCCTCCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTA 1200

Qy 1201 TGGTCTTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTTACCCCCATGGCAGAGACCC 1260
 |||

Db 1201 TGGTCTTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTTACCCCCATGGCAGAGACCC 1260

Qy 1261 ACCGCTGTTCTGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGA 1320
 |||

Db 1261 ACCGCTGTTCTGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGA 1320

Qy 1321 CCTTCGATAAAATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAACAGCACTTGTC 1380
 |||

Db 1321 CCTTCGATAAAATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAACAGCACTTGTC 1380

Qy 1381 ACCA 1384
 |||

Db 1381 ACCA 1384

RESULT 2

AB010123

LOCUS AB010123 2104 bp mRNA linear ROD 06-DEC-2001

DEFINITION Mus musculus mRNA for AMSH, complete cds.

ACCESSION AB010123

VERSION AB010123.1 GI:17385633

KEYWORDS AMSH.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1
 AUTHORS Ishii,N., Owada,Y., Yamada,M., Miura,S., Murata,K., Asao,H.,
 Kondo,H. and Sugamura,K.

TITLE Loss of neurons in the hippocampus and cerebral cortex of
 AMSH-deficient mice

JOURNAL Mol. Cell. Biol. 21 (24), 8626-8637 (2001)

MEDLINE 21570416

PUBMED 11713295

REFERENCE

2 (bases 1 to 2104)
 AUTHORS Ishii,N., Aizawa,K. and Sugamura,K.

TITLE Direct Submission

JOURNAL Submitted (06-JAN-1998) Kazuo Sugamura, Tohoku University School of
 Medicine, Department of Microbiology and Immunology; 2-1
 Seiryomachi, Aoba-ku, Sendai, Miyagi 980-8575, Japan
 (E-mail:sugamura@mail.cc.tohoku.ac.jp, Tel:81-22-717-8096,
 Fax:81-22-717-8097)

FEATURES

source

Location/Qualifiers

1. .2104

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/tissue_type="brain"

Db	481	GGAAAAAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATT	540
Qy	541	CCATGCCTTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGT	600
Db	541	CCATGCCTTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGT	600
Qy	601	TCAAGAGTTCGGGAAGGTAGACCCTGGCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAA	660
Db	601	TCAAGAGTTCGGGAAGGTAGACCCTGGCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAA	660
Qy	661	GCCTTGTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAA	720
Db	661	GCCTTGTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAA	720
Qy	721	CACAGGCATGAGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTT	780
Db	721	CACAGGCATGAGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTT	780
Qy	781	AAGCGTCATAGAAAATGTTCCCACCATTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAA	840
Db	781	AAGCGTCATAGAAAATGTTCCCACCATTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAA	840
Qy	841	TCTGTGCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTG	900
Db	841	TCTGTGCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTG	900
Qy	901	TGGAGTCCTCTGTGGAAAACCTGATGAGAAATGAATTCACAATCACACATGTTCTCATCCC	960
Db	901	TGGAGTCCTCTGTGGAAAACCTGATGAGAAATGAATTCACAATCACACATGTTCTCATCCC	960
Qy	961	CAGACAAAATGGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTAT	1020
Db	961	CAGACAAAATGGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTAT	1020
Qy	1021	GCAGGATGACCTTGGACTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGC	1080
Db	1021	GCAGGATGACCTTGGACTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGC	1080
Qy	1081	CTTTCTGTCCAGTGTGGATCTCCCACTCACTGCTCCTACCAAATGATGTTACCAGAGTC	1140
Db	1081	CTTTCTGTCCAGTGTGGATCTCCCACTCACTGCTCCTACCAAATGATGTTACCAGAGTC	1140
Qy	1141	CATCGCAATCGTCTGTTCCCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTA	1200
Db	1141	CATCGCAATCGTCTGTTCCCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTA	1200
Qy	1201	TGGTCTTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTCACCCCCATGGCAGAGACCC	1260
Db	1201	TGGTCTTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTCACCCCCATGGCAGAGACCC	1260
Qy	1261	ACCGCTGTTCTGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGA	1320
Db	1261	ACCGCTGTTCTGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGA	1320
Qy	1321	CCTTCGATAAATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAAACAGCACTTGTC	1380
Db	1321	CCTTCGATAAATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAAACAGCACTTGTC	1380

Qy 1381 ACCA 1384
 ||||
Db 1381 ACCA 1384

RESULT 3

BC025111

LOCUS BC025111 2094 bp mRNA linear ROD 12-NOV-2003

DEFINITION Mus musculus Stam binding protein, mRNA (cDNA clone MGC:36212
IMAGE:4194332), complete cds.

ACCESSION BC025111

VERSION BC025111.1 GI:19263555

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2094)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 2094)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 61 Row: j Column: 3

This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 17941276.

FEATURES

	Location/Qualifiers
source	1. .2094 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:36212 IMAGE:4194332" /tissue_type="Liver, normal. 5 month old male mouse." /clone_lib="NCI CGAP_Li9" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"
gene	1. .2094 /gene="Stambp" /note="synonym: Amsh" /db_xref="LocusID:70527" /db_xref="MGI:1917777"
CDS	51. .1325 /codon_start=1 /product="Stam binding protein" /protein_id="AAH25111.1" /db_xref="GI:19263556" /db_xref="LocusID:70527" /translation="MSDHGDVSLPPQDRVRIILSQLGSAVELNEDI PPRRYRSGVEII RMASVYSEEGNIEHAFILYNKYITLFIKLPKHRDYKSAI IPEKKDAVKKLKSVAFPK AEELKTELLRRYTKEYEQYKERKKKEEEELARNIAIQQELEKEKQ RVAQQKQKQLEQE QFHAFEEIMIQRQELEKERLKI VQEFQKVDGPGCGPLLPDLEKPCVDVAPSSPFSPTQT PDCNTGMRPAKPPVDRSLKPGALSVIENVPTIEGLRHIVVPRNLCSEFLQLASANTA KGIETCGVLCGLMRNEFTITHVLI PRQNGGPDYCHTENEEEIFFMQDDLGLLTLGWI HTHPTQTAFLLSSVDLHHCYSYQMMLPESIAIVCSPKFQETGFFKLT DYGLQEISTCRQ KGFHPHGRDPPLFCDCSHVTVKDRIVTITDLR"
misc_feature	804. .1133 /note="Mov34; Region: Mov34/MPN/PAD-1 family. Members of this family are found in proteasome regulatory subunits, eukaryotic initiation factor 3 (eIF3) subunits and regulators of transcription factors. This family is also known as the MPN domain and PAD-1-like domain. It has been shown that this domain occurs in prokaryotes" /db_xref="CDD:pfam01398"

ORIGIN

Query Match 99.6%; Score 1379; DB 10; Length 2094;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GTTTCGGAAGCTCTGACTGTCATCCTTCACGAAAGAACTTATTTGTCCAATGTCTGACC 65

Db	1		1	GTTCGGAAGCTCTGACTGTCATCCTTCACGAAAGAACTTATTTGTCCAATGTCTGACC	60
Qy	66		66	ATGGGGATGTGAGCCTCCACCCCAAGACCGGGTGAGGATTCTGTCCCAACTGGGAGTG	125
Db	61		61	ATGGGGATGTGAGCCTCCACCCCAAGACCGGGTGAGGATTCTGTCCCAACTGGGAGTG	120
Qy	126		126	CAGTTGAGTTAAATGAAGACATTCCACCCCGTCGCTACTACCGCTCCGGTGTTGAGATCA	185
Db	121		121	CAGTTGAGTTAAATGAAGACATTCCACCCCGTCGCTACTACCGCTCCGGTGTTGAGATCA	180
Qy	186		186	TCCGCATGGCGTCCGTTTACTCGGAAGAAGGCAACATTGAACATGCCTTTATCCTCTACA	245
Db	181		181	TCCGCATGGCGTCCGTTTACTCGGAAGAAGGCAACATTGAACATGCCTTTATCCTCTACA	240
Qy	246		246	ACAAGTACATCACGCTGTTTATTGAAAAACTTCCGAAACACCGAGACTACAAATCAGCTA	305
Db	241		241	ACAAGTACATCACGCTGTTTATTGAAAAACTTCCGAAACACCGAGACTACAAATCAGCTA	300
Qy	306		306	TCATTCTGAGAAGAAAGATGCTGTCAAGAAATTAAAGAGCGTCGCTTTCCCTAAAGCGG	365
Db	301		301	TCATTCTGAGAAGAAAGATGCTGTCAAGAAATTAAAGAGCGTCGCTTTCCCTAAAGCGG	360
Qy	366		366	AAGAGCTGAAGACAGAGCTCTTGAGAAGATACACCAAAGAATATGAGCAGTATAAAGAGC	425
Db	361		361	AAGAGCTGAAGACAGAGCTCTTGAGAAGATACACCAAAGAATATGAGCAGTATAAAGAGC	420
Qy	426		426	GAAAGAAAAAGGAAGAAGAGGAACTTGCCCGAAATATCGCCATCCAGCAAGAGTTGGAAA	485
Db	421		421	GAAAGAAAAAGGAAGAAGAGGAACTTGCCCGAAATATCGCCATCCAGCAAGAGTTGGAAA	480
Qy	486		486	AAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATTCCATG	545
Db	481		481	AAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATTCCATG	540
Qy	546		546	CCTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGTTCAAG	605
Db	541		541	CCTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGTTCAAG	600
Qy	606		606	AGTTCGGGAAGGTAGACCCTGGCCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAAGCCTT	665
Db	601		601	AGTTCGGGAAGGTAGACCCTGGCCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAAGCCTT	660
Qy	666		666	GTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAACACAG	725
Db	661		661	GTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAACACAG	720
Qy	726		726	GCA TGAGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTTAAGCG	785
Db	721		721	GCA TGAGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTTAAGCG	780
Qy	786		786	TCATAGAAAAATGTTCCCAACCATTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAATCTGT	845
Db	781		781	TCATAGAAAAATGTTCCCAACCATTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAATCTGT	840
Qy	846		846	GCTCAGAAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTGTGGAG	905

Db 841 GCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTGTGGAG 900

QY 906 TCCTCTGTGGAAAAGTCTGATGAGAAATGAATTCACAATCACACATGTTCTCATCCCCAGAC 965
 |||

Db 901 TCCTCTGTGGAAAAGTCTGATGAGAAATGAATTCACAATCACACATGTTCTCATCCCCAGAC 960

QY 966 AAAATGGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTCTTTATGCAGG 1025
 |||

Db 961 AAAATGGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTCTTTATGCAGG 1020

QY 1026 ATGACCTTGGACTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGCCTTTC 1085
 |||

Db 1021 ATGACCTTGGACTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGCCTTTC 1080

QY 1086 TGTCCAGTGTGGATCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTCCATCG 1145
 |||

Db 1081 TGTCCAGTGTGGATCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTCCATCG 1140

QY 1146 CAATCGTCTGTTCCCCAAAGTTCCAGGAAAGTGGATTCTTTAAGCTAACTGACTATGGTC 1205
 |||

Db 1141 CAATCGTCTGTTCCCCAAAGTTCCAGGAAAGTGGATTCTTTAAGCTAACTGACTATGGTC 1200

QY 1206 TTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTCACCCCCATGGCAGAGACCCACCGC 1265
 |||

Db 1201 TTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTCACCCCCATGGCAGAGACCCACCGC 1260

QY 1266 TGTCTGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGACCTTC 1325
 |||

Db 1261 TGTCTGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGACCTTC 1320

QY 1326 GATAAATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAACAGCACTTGTACCA 1384
 |||

Db 1321 GATAAATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAACAGCACTTGTACCA 1379

RESULT 4

BC003497

LOCUS BC003497 2122 bp mRNA linear ROD 03-OCT-2003
 DEFINITION Mus musculus Stam binding protein, mRNA (cDNA clone MGC:6953
 IMAGE:3153829), complete cds.

ACCESSION BC003497

VERSION BC003497.1 GI:13097542

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2122)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 2122)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 5 Row: p Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 17941276.

FEATURES

source Location/Qualifiers

1. .2122
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="MGC:6953 IMAGE:3153829"
/tissue_type="Mammary tumor. Brca1-/fl; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI_CGAP_Mam3"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1. .2122
/gene="Stambp"
/note="synonym: Amsh"
/db_xref="LocusID:70527"

CDS

```
/db_xref="MGI:1917777"
52. .1326
/codon_start=1
/product="Stam binding protein"
/protein_id="AAH03497.1"
/db_xref="GI:13097543"
/db_xref="LocusID:70527"
/translation="MSDHGDVSLPPQDRVRILSQLGSAVELNEDIPPRRYYRSGVEI I
RMASVYSEEGNIEHAFILYNKYITLFIKLPKHRDYKSAI IPEKKDAVKKLKSVAFPK
AEELKTELLRRYTKEYEQYKERKKKEEEELARNIAIQQELEKEKQ RVAQQKQKQLEQE
QFHAFEEMIQRQELEKERLKI VQEF GKVDPGPCG PLLPDLEKPCVDVAPSSPFSPTQT
PDCNTGMRPAKPPVVDRLKPGALSVIENVPTIEGLRHIVVPRNLCSEFLQLASANTA
KGIETCGVLCGKLMRNEFTITHVLI PRQNGGPDYCHTENEEEIFFMQDDLGLLTGWI
HTHTPTQTAFLLSSVDLHTHCSYQMMLPESIAIVCSPK FQETGFFKLT DYGLQEISTCRQ
KGFHPHGRDPPLFCDCSHVTVKDRIVTITDLR"
```

misc_feature

```
805. .1134
/note="Mov34; Region: Mov34/MPN/PAD-1 family. Members of
this family are found in proteasome regulatory subunits,
eukaryotic initiation factor 3 (eIF3) subunits and
regulators of transcription factors. This family is also
known as the MPN domain and PAD-1-like domain. It has been
shown that this domain occurs in prokaryotes"
/db_xref="CDD:pfam01398"
```

ORIGIN

```
Query Match          98.8%;  Score 1367.6;  DB 10;  Length 2122;
Best Local Similarity 99.7%;  Pred. No. 0;
Matches 1370;  Conservative 0;  Mismatches 4;  Indels 0;  Gaps 0;
```

```
Qy      11  CGGAAGCTCTGACTGTCTATCCTTCACGAAAGAACTTATTTGTCCAATGTCTGACCATGGG 70
      ||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7  CGTCCGCTCTGACTGTCTATCCTTCACGAAAGAACTTATTTGTCCAATGTCTGACCATGGG 66

Qy     71  GATGTGAGCCTCCCACCCCAAGACCGGGTGAGGATTCTGTCCCAACTTGGGAGTGCAGTT 130
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67  GATGTGAGCCTCCCACCCCAAGACCGGGTGAGGATTCTGTCCCAACTTGGGAGTGCAGTT 126

Qy    131  GAGTTAAATGAAGACATTCACCCCGTCGCTACTACCGCTCCGGTGTTGAGATCATCCGC 190
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127  GAGTTAAATGAAGACATTCACCCCGTCGCTACTACCGCTCCGGTGTTGAGATCATCCGC 186

Qy    191  ATGGCGTCCGTTTACTCGGAAGAAGGCAACATTGAACATGCCTTTATCCTCTACAACAAG 250
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187  ATGGCGTCCGTTTACTCGGAAGAAGGCAACATTGAACATGCCTTTATCCTCTACAACAAG 246

Qy    251  TACATCACGCTGTTTATTGAAAAACTTCCGAAACACCGAGACTACAAATCAGCTATCATT 310
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247  TACATCACGCTGTTTATTGAAAAACTTCCGAAACACCGAGACTACAAATCAGCTATCATT 306

Qy    311  CCTGAGAAGAAAGATGCTGTCAAGAAATTAAAGAGCGTCGCTTTCCCTAAAGCGGAAGAG 370
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307  CCTGAGAAGAAAGATGCCGTCAAGAAATTAAAGAGCGTCGCTTTCCCTAAAGCGGAAGAG 366

Qy    371  CTGAAGACAGAGCTCTTGAGAAGATACACCAAAGAATATGAGCAGTATAAAGAGCGGAAAG 430
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367  CTGAAGACAGAGCTCTTGAGAAGATACACCAAAGAATATGAGCAGTATAAAGAGCGGAAAG 426
```


Qy	431	AAAAAGGAAGAAGAGGAACTTGCCCGAAATATCGCCATCCAGCAAGAGTTGGAAAAAGAA	490
Db	427	AAAAAGGAAGAAGAGGAACTTGCCCGAAATATCGCCATCCAGCAAGAGTTGGAAAAAGAA	486
Qy	491	AAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATTCCATGCCTTT	550
Db	487	AAACAGAGGGTTGCTCAGCAGAAGCAGAAGCAGCTAGAGCAGGAGCAATTCCATGCCTTT	546
Qy	551	GAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGTTCAAGAGTTC	610
Db	547	GAGGAGATGATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGTTCAAGAGTTC	606
Qy	611	GGGAAGGTAGACCCTGGCCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAAGCCTTGTGTA	670
Db	607	GGGAAGGTAGACCCTGGCCCCCTGCGGGCCTCTGCTCCCTGATCTGGAAAAGCCTTGTGTA	666
Qy	671	GATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAACACAGGCATG	730
Db	667	GATGTGGCCCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAACACAGGCATG	726
Qy	731	AGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTTAAGCGTCATA	790
Db	727	AGGCCAGCTAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTTAAGCGTCATA	786
Qy	791	GAAAATGTTCCACCATTTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAATCTGTGCTCA	850
Db	787	GAAAATGTTCCACCATTTGAAGGCCTGCGCCACATCGTGGTGCCCCGTAATCTGTGCTCA	846
Qy	851	GAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTGTGGAGTCCTC	910
Db	847	GAATTTCTCCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTGTGGAGTCCTC	906
Qy	911	TGTGGAAAACCTGATGAGAAATGAATTCACAATCACACATGTTCTCATCCCCAGACAAAAT	970
Db	907	TGTGGAAAACCTGATGAGAAATGAATTCACAATCACACATGTTCTCATCCCCAGACAAAAT	966
Qy	971	GGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTATGCAGGATGAC	1030
Db	967	GGTGGGCCTGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTATGCAGGATGAC	1026
Qy	1031	CTTGGACTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGCCTTTCTGTCC	1090
Db	1027	CTTGGACTCCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGCCTTTCTGTCC	1086
Qy	1091	AGTGTGGATCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTCCATCGCAATC	1150
Db	1087	AGTGTGGATCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTCCATCGCAATC	1146
Qy	1151	GTCTGTTCCCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTATGGTCTTCAA	1210
Db	1147	GTCTGTTCCCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTATGGTCTTCAA	1206
Qy	1211	GAGATTTCAACCTGCCGGCAGAAAGGCTTTTACCCCCATGGCAGAGACCCACCGCTGTTC	1270
Db	1207	GAGATTTCAACCTGCCGGCAGAAAGGCTTTTACCCCCATGGCAGAGACCCACCGCTGTTC	1266
Qy	1271	TGTGACTGCAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGACCTTCGATAA	1330

Db 1267 TGTGACTGCAGCCATGTCACCTGTCAAGGACAGAATTGTGACGATCACAGACCTTCGATAA 1326
 Qy 1331 ATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAACAGCACTTGTCAACCA 1384
 Db 1327 ATCTCAAATCATGAACCAGGGAGATGGATCACTGGGTAACAGCACTTGTCAACCA 1380

RESULT 5

BC006939

LOCUS BC006939 2172 bp mRNA linear ROD 03-OCT-2003

DEFINITION Mus musculus Stam binding protein, mRNA (cDNA clone MGC:6839 IMAGE:2649718), complete cds.

ACCESSION BC006939

VERSION BC006939.1 GI:13905279

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2172)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 2172)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (27-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 5 Row: h Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES	Location/Qualifiers
source	1. .2172 /organism="Mus musculus" /mol_type="mRNA" /strain="129,C57BL/6J,FVB/N" /db_xref="taxon:10090" /clone="MGC:6839 IMAGE:2649718" /tissue_type="Mammary tumor. Brcal-/fl; MMTV-Cre model. 10 months old, gross tissue." /clone_lib="NCI_CGAP_Mam3" /lab_host="DH10B"
gene	/note="Vector: pCMV-SPORT6" 1. .2172 /gene="Stamp" /note="synonym: Amsh" /db_xref="LocusID:70527" /db_xref="MGI:1917777"
CDS	107. .1381 /codon_start=1 /product="Stamp protein" /protein_id="AAH06939.1" /db_xref="GI:13905280" /db_xref="LocusID:70527" /translation="MSDHGDVSLPPQDRVRILSQLGSAVELNEDI PPRYYRSGVEII RMASVYSEEGNIEHAFILYNKYITLFIKLPKHRDYKSAI IPEKKDAVKKLKSVAFPK AEELKTELLRRYTKEYEQYKERKKKEEEELARNIAIQQELEKEKQVAQQKQKQLEQE QFHAFEEMIQRQELEKERLKI VQEFQKVDPGPCGPLLDPLEKPCVDVAPSSPFSPTQT PDCNTGMRPAKPPVDRSLKPGALSVIENVPTIEGLRHIVVPRNLCSEFLQLASANTA KGIETCGVLCGKLMRNEFTITHVLI PRQNGGPDYCHTENEEEIFFMQDDLGLLTLGWI HTHPTQTAFLLSSVDLHTHCSYQMMLPESIAIVCSPKQFQETGFFKLTDYGLQEISTCRQ KGFHPHGRDPPLFCDCSHVTVKDRIVTITDLR"
misc_feature	860. .1189 /note="Mov34; Region: Mov34/MPN/PAD-1 family. Members of this family are found in proteasome regulatory subunits, eukaryotic initiation factor 3 (eIF3) subunits and regulators of transcription factors. This family is also known as the MPN domain and PAD-1-like domain. It has been shown that this domain occurs in prokaryotes" /db_xref="CDD:pfam01398"

ORIGIN

Query Match 97.0%; Score 1343; DB 10; Length 2172;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	42	AACTTATTTGTCCAATGTCTGACCATGGGGATGTGAGCCTCCCACCCCAAGACCGGGTGA	101
Db	93	AACTTATTTGTCCAATGTCTGACCATGGGGATGTGAGCCTCCCACCCCAAGACCGGGTGA	152
Qy	102	GGATTCTGTCCCAACTTGGGAGTGAGTTGAGTTAAATGAAGACATTCCACCCCGTCGCT	161
Db	153	GGATTCTGTCCCAACTTGGGAGTGAGTTGAGTTAAATGAAGACATTCCACCCCGTCGCT	212
Qy	162	ACTACCGCTCCGGTGTTGAGATCATCCGCATGGCGTCCGTTTACTCGGAAGAAGGCAACA	221
Db	213	ACTACCGCTCCGGTGTTGAGATCATCCGCATGGCGTCCGTTTACTCGGAAGAAGGCAACA	272
Qy	222	TTGAACATGCCTTTATCCTCTACAACAAGTACATCAGCTGTTTATTGAAAACTTCCGA	281
Db	273	TTGAACATGCCTTTATCCTCTACAACAAGTACATCAGCTGTTTATTGAAAACTTCCGA	332
Qy	282	AACACCGAGACTACAAATCAGCTATCATTCCTGAGAAGAAAGATGCTGTCAAGAAATTAA	341
Db	333	AACACCGAGACTACAAATCAGCTATCATTCCTGAGAAGAAAGATGCTGTCAAGAAATTAA	392
Qy	342	AGAGCGTCGCTTTCCTAAAGCGGAAGAGCTGAAGACAGAGCTCTTGAGAAGATACACCA	401
Db	393	AGAGCGTCGCTTTCCTAAAGCGGAAGAGCTGAAGACAGAGCTCTTGAGAAGATACACCA	452
Qy	402	AAGAATATGAGCAGTATAAAGAGCGAAAGAAAAAGGAAGAAGAGGAACTTGCCCGAAATA	461
Db	453	AAGAATATGAGCAGTATAAAGAGCGAAAGAAAAAGGAAGAAGAGGAACTTGCCCGAAATA	512
Qy	462	TCGCCATCCAGCAAGAGTTGGAAGAAAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGC	521
Db	513	TCGCCATCCAGCAAGAGTTGGAAGAAAGAAAAACAGAGGGTTGCTCAGCAGAAGCAGAAGC	572
Qy	522	AGCTAGAGCAGGAGCAATTCCATGCCTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAA	581
Db	573	AGCTAGAGCAGGAGCAATTCCATGCCTTTGAGGAGATGATCCAGAGGCAGGAGCTGGAAA	632
Qy	582	AAGAACGGCTGAAAATTGTTCAAGAGTTCGGGAAGGTAGACCTGGCCCCCTGCGGGCCTC	641
Db	633	AAGAACGGCTGAAAATTGTTCAAGAGTTCGGGAAGGTAGACCTGGCCCCCTGCGGGCCTC	692
Qy	642	TGCTCCCTGATCTGGAAGAGCCTTGTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCCA	701
Db	693	TGCTCCCTGATCTGGAAGAGCCTTGTGTAGATGTGGCCCCCAGCTCACCGTTCTCGCCCA	752
Qy	702	CGCAGACTCCAGACTGTAACACAGGCATGAGGCCAGCTAAGCCACCTGTGGTGACAGGT	761
Db	753	CGCAGACTCCAGACTGTAACACAGGCATGAGGCCAGCTAAGCCACCTGTGGTGACAGGT	812
Qy	762	CCCTGAAACCTGGAGCGTTAAGCGTCATAGAAAATGTTCCACCATTTGAAGGCCTGCGCC	821
Db	813	CCCTGAAACCTGGAGCGTTAAGCGTCATAGAAAATGTTCCACCATTTGAAGGCCTGCGCC	872
Qy	822	ACATCGTGGTGCCCCGTAATCTGTGCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCG	881

Db 873 ACATCGTGGTGCCCCGTAATCTGTGCTCAGAATTTCTCCAGCTTGCCAGTGCCAATACCG 932
 QY 882 CCAAAGGCATTGAAACCTGTGGAGTCTCTGTGGAAAACCTGATGAGAAATGAATTCACAA 941
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 933 CCAAAGGCATTGAAACCTGTGGAGTCTCTGTGGAAAACCTGATGAGAAATGAATTCACAA 992
 QY 942 TCACACATGTTCTCATCCCCAGACAAAATGGTGGGCCTGATTATTGCCACACGGAGAATG 1001
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 993 TCACACATGTTCTCATCCCCAGACAAAATGGTGGGCCTGATTATTGCCACACGGAGAATG 1052
 QY 1002 AAGAAGAAATTTTCTTTATGCAGGATGACCTTGGACTCCTCACTCTTGGCTGGATCCATA 1061
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1053 AAGAAGAAATTTTCTTTATGCAGGATGACCTTGGACTCCTCACTCTTGGCTGGATCCATA 1112
 QY 1062 CTCATCCAACCCAAACGGCCTTTCTGTCCAGTGTGGATCTCCACACTCACTGCTCCTACC 1121
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1113 CTCATCCAACCCAAACGGCCTTTCTGTCCAGTGTGGATCTCCACACTCACTGCTCCTACC 1172
 QY 1122 AAATGATGTTACCAGAGTCCATCGCAATCGTCTGTTCCCCAAAGTTCCAGGAAACTGGAT 1181
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1173 AAATGATGTTACCAGAGTCCATCGCAATCGTCTGTTCCCCAAAGTTCCAGGAAACTGGAT 1232
 QY 1182 TCTTTAAGCTAACTGACTATGGTCTTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTC 1241
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1233 TCTTTAAGCTAACTGACTATGGTCTTCAAGAGATTTCAACCTGCCGGCAGAAAGGCTTTC 1292
 QY 1242 ACCCCCATGGCAGAGACCCACCGCTGTTCTGTGACTGCAGCCATGTCACTGTCAAGGACA 1301
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1293 ACCCCCATGGCAGAGACCCACCGCTGTTCTGTGACTGCAGCCATGTCACTGTCAAGGACA 1352
 QY 1302 GAATTGTGACGATCACAGACCTTCGATAAATCTCAAATCATGAACCAGGGAGATGGATCA 1361
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1353 GAATTGTGACGATCACAGACCTTCGATAAATCTCAAATCATGAACCAGGGAGATGGATCA 1412
 QY 1362 CTGGGTAACAGCACTTGTACCA 1384
 ||||||||||||||||||
 Db 1413 CTGGGTAACAGCACTTGTACCA 1435

RESULT 6

AY083159

LOCUS AY083159 1544 bp mRNA linear ROD 26-MAR-2002

DEFINITION Rattus norvegicus AMSH mRNA, complete cds.

ACCESSION AY083159

VERSION AY083159.1 GI:19743767

KEYWORDS

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 1544)

AUTHORS Pawlak,A. and Guellaen,G.

TITLE Direct Submission

JOURNAL Submitted (08-MAR-2002) U99, INSERM, Hopital Henri Mondor, Creteil
94010, France

FEATURES Location/Qualifiers

Db	494	AGTTGCACAGCAGAAGCAGAAGCAGCTCGAGCAGGAGCAGTTCCATGCCTTTGAGAAGAT	553
Qy	559	GATCCAGAGGCAGGAGCTGGAAAAAGAACGGCTGAAAATTGTTCAAGAGTTCGGGAAGGT	618
Db	554	GATCCAGAAGCAGGAGCTAGAGAAAGAGCGGCTAAAAATTGTTCAAGAGTTCGGGAAGGT	613
Qy	619	AGACCCCTGGCCCCGTCGGGCCCTCTGCTCCCTGATCTGGAAGCCTTGTGTAGATGTGGC	678
Db	614	AGACCCCTGGCCCCGTCGGGCCCTCTGCTCCCTGATCTGGAAGCCTTGTGTAGATGTGGC	673
Qy	679	CCCCAGCTCACCGTTCTCGCCACGCAGACTCCAGACTGTAACACAGGCATGAGGCCAGC	738
Db	674	CCCCAGTTCACCTTTCTCGCCACGCAGACTTCAGACTGTAACACAACCTTGAGGCCAGC	733
Qy	739	TAAGCCACCTGTGGTGGACAGGTCCCTGAAACCTGGAGCGTTAAGCGTCATAGAAAATGT	798
Db	734	TAAGCCACCTGTGGTGGACAGGTCCCTCAAACCTGGAGCATTAAGCGTCATAGAAAATGT	793
Qy	799	TCCCACCATTGAAGGCCTGCGCCACATCGTGGTGCCCGTAATCTGTGCTCAGAAATTTCT	858
Db	794	TCCCACCATTGAAGGCCTGCGCCACATTTGTGGTGCCCGCAATCTGTGCTCAGAAATTTCT	853
Qy	859	CCAGCTTGCCAGTGCCAATACCGCCAAAGGCATTGAAACCTGTGGAGTCCCTCTGTGGAAA	918
Db	854	CCAGCTTGCCAGCGCCAACACTGCCAAAGGCATCGAGACCTGTGGAGTCCCTCTGTGGAAA	913
Qy	919	ACTGATGAGAAAATGAATTCACAATCACACATGTTCTCATCCCCAGACAAAATGGTGGGCC	978
Db	914	ACTGATGAGAAAATGAATTCACAATCACACATGTTCTCATCCCCAGACAAAATGGTGGGCC	973
Qy	979	TGATTATTGCCACACGGAGAATGAAGAAGAAATTTTCTTTATGCAGGATGACCTTGGACT	1038
Db	974	TGATTATTGCCACACAGAGAATGAAGAAGAAATTTTCTTTATGCAGGATGATCTTGGACT	1033
Qy	1039	CCTCACTCTTGGCTGGATCCATACTCATCCAACCCAAACGGCCTTTCTGTCCAGTGTGGA	1098
Db	1034	CCTCACTCTTGGCTGGATCCACACCCATCCAACCCAAACGGCCTTTCTGTCCAGTGTGGA	1093
Qy	1099	TCTCCACACTCACTGCTCCTACCAAATGATGTTACCAGAGTCCATCGCAATCGTCTGTTT	1158
Db	1094	TCTGCACACGCACTGCTCCTACCAAATGATGTTACCAGAGTCCATAGCAATTGTCTGCTC	1153
Qy	1159	CCCAAAGTTCCAGGAAACTGGATTCTTTAAGCTAACTGACTATGGTCTTCAAGAGATTTT	1218
Db	1154	CCCCAAGTTCCAGGAGACTGGATTCTTTAAATTAAGTAACTGACTATGGCCTTCAAGAGATTT	1213
Qy	1219	AACCTGCCGGCAGAAAGGCTTTACCCCCATGGCAGAGACCCACCGCTGTTCTGTGACTG	1278
Db	1214	AACCTGCCGGCAGAAAGGCTTTACCCCCATGGCAGAGACCCACCGCTGTTCTGTGACTG	1273
Qy	1279	CAGCCATGTCACTGTCAAGGACAGAATTGTGACGATCACAGACCTTCGATAAATCTCAA	1338
Db	1274	CAGCCATGTCACTGTCAAAGACAGAATTGTGACGATCACAGACCTTCGATAAATCTC-AG	1332
Qy	1339	TCATGAACCAGGGAGATGGATCACTGGGTAAACAGCACTTGTCACCA	1384
Db	1333	TCATGAACCA-GGAGGTGGCCCACTGGGTAAACACACTTGCCACCA	1377